DOCUMENT RESUME

ED 124 921 CS 002 768

AUTHOR Anderson, Linda Mahaffey; Brophy, Jere E.
TITLE An Experimental Investigation of First Grade Reading Group Instruction. Report No. 76-3.
INSTITUTION Texas Univ., Austin. Research and Development Center for Teacher Education.
PUB DATE 76
NOTE 24p.; Report presented at the Annual Meeting of the American Educational Research Association (San Francisco, California, April 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Ability Grouping; Beginning Reading; *Effective Teaching; Grade 1; *Grouping (Instructional Purposes); Homogeneous Grouping; Primary Education; *Reading Instruction; Reading Research; *Small Group Instruction; *Teacher Behavior; Teaching Methods

ABSTRACT

The purpose of this study was to determine the effectiveness of 22 principles of small-group instruction of young children. The principles dealt only with teacher behavior in managing instruction in the small-group setting and were organized in two major sections. Items described as organization and management concerned getting the children's attention, introducing the lesson, calling on children, and meeting individual learning needs within the groups. Items included under the second category, responding to children's answers, suggested appropriate teacher behavior when a child does not respond, when the child's answer is incorrect, and when the child is correct, and commented on praise and criticism. In the experiment, teachers in nine schools were divided into three groups: treatment-observed, in which ten teachers were instructed in the 22 principles and were observed once a week for most of the year; treatment-unobserved, in which seven teachers were instructed in the 22 principles but were not subsequently observed; and control-observed. The results indicated that the treatment group teachers produced greater reading gains in the students than did the control group teachers. (MKM)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED124921

An Experimental Investigation of First Grade Reading Group Instruction[1]

Linda Mahaffey Anderson

and

Jere E. Brophy

Report No. 76-3

Research and Development Center for Teacher Education

The University of Texas at Austin

April, 1976

[1]Paper presented at the annual meeting of the American Educational Research Association, San Francisco, California, April, 1976.

CS002768

This project was supported by the National Institute of Education Contract OEC 6-10-108, Research and Development Center for Teacher Education, and by Contract NIE-C-74-0089, Correlates of Effective Teaching. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by that office should be inferred.

This paper is the first in a series of reports about the First Grade Reading Group Study. As further analyses are completed, future reports will expand on these findings and relate them to other data. The results presented here are the comparisons of reading scores made in treatment and control classrooms. Future reports will discuss the relationships between these gains and implementation data based on classroom observations.

Background of the Study

The purpose of the First Grade Reading Group Study was to determine the effectiveness of 22 principles of small-group instruction of young children. These principles were selected on the basis of previous correlational research (Brophy & Evertson, 1976) and program development efforts (Southwest Educational Development Laboratory, 1973) which suggested promising approaches to teaching young children in small group settings. Since much of the previous research was correlational in nature, the First Grade Reading Group Study was an attempt to verify the usefulness of these principles in an experimental situation.

The first grade reading group setting was picked because of the age of the children (previous research dealt with preschool and early elementary teaching) and the fact that this instructional format is nearly universal at this grade level. Reading is taught in all first grade classrooms, and, although there have been recent trends toward individualized instruction, most first grade reading instruction still is offered in the context of the ability-based small group. Therefore, findings from this study will be useful to the large body of teachers who use this instructional format in teaching beginning reading.

The study was a direct outgrowth of The Texas Teacher Effectiveness Study (Brophy & Evertson, 1976). These findings of this earlier study were based on classroom observations of second and third grade teachers who varied in their teaching effectiveness as measured by residualized gain scores on achievement tests given to their students annually over a four-year period. These teachers' behaviors were correlated with their effectiveness scores to identify teaching behaviors which were associated with student learning gains. The results suggested several directions for future experimental research, and the present study was an effort to examine some of these behaviors within an experimental framework. The correlational findings could not be interpreted as causative in nature (i.e., one could not conclude that the teaching behaviors caused learning gains in the students), although the implications were strong. However, an intervention study involving encouraging teachers to do certain things and then measuring the relationships of those behaviors to learning would be a much stronger indication of causal relationships between teaching behaviors and student outcomes.

Analysis of the correlational research relating to small group teaching led to the development of an instructional model consisting of 22 principles. Analyses of the data collected in this study will examine the effects of the model as a whole as well as of its 22 components.

The 22 principles deal only with teacher behavior in managing instruction in the small group setting. Thus, they are "curriculum free." They could be adapted to any reading series or instructional content chosen by the teacher, so long as instruction was conducted within the traditional small group context. The teacher could follow these techniques and also conform

to the suggestions offered in the teachers' manuals about sequencing of instruction, questions to be asked, follow-up activities, and so on.

The principles were organized into two major sections: (1) organization and management, and (2) responding to individual differences in the group setting. A summary of the principles is presented below.

I. ORGANIZATION AND MANAGEMENT

GETTING THE CHILDREN'S ATTENTION

1. The teacher gets everyone's attention before starting the lesson.
2. The children sit with their backs to the rest of the class while the teacher faces the class.

INTRODUCING THE LESSON

3. The teacher introduces the lesson with a brief overview.
4. The teacher presents new words clearly.
5. After presenting new words, the teacher has the children repeat them.
6. A demonstration or explanation precedes the children's attempts to do the work.

CALLING ON CHILDREN

7. The teacher should work with one child at a time, so that everyone is checked and receives feedback.
8. The teacher should call on children in order rather than randomly.
9. Occasionally the teacher should question a child about another child's responses (to keep everyone alert).

10. The teacher should minimize calling on volunteers.

11. The teacher should discourage call outs and should emphasize that each child is responsible for the question asked of him.

12. The teacher should avoid rhetorical questions, answering her own questions, or repeating questions. These confuse the children.

MEETING INDIVIDUAL LEARNING NEEDS WITHIN THE GROUP

13. At some point, the teacher must decide if the whole group can meet the lesson's objectives. If she decides they can, she should hold the group together, making sure that everyone masters each step before moving on to the next step.

14. If the teacher decides that everyone cannot meet the objective, the students who can do so should be taught through to the end and then dismissed, so that the teacher can spend more time with the other children.

15. An exception to the above occurs when the teacher wants to use a student who has mastered the objective as a model for the others. Here, she may retain one or more such students in the group in order to carry on a dialogue.

16. If some of the children do not succeed in meeting the objectives before lesson time is up, arrangements should be made for extra tutorial help.

II. RESPONDING TO CHILDREN'S ANSWERS

The teacher's feedback to children's answers depends on (1) the type of question (whether it requires memory or reasoning), (2) the pace of questioning (whether rapid for drill or slower for more thoughtful questions),

and (3) the child's answer (correct, incorrect, "I don't know," or no response).

WHEN THE CHILD DOES NOT RESPOND

17. After asking a question, the teacher waits for the child to respond and also sees that other children wait and do not call out answers. During rapid pacing, she waits a few seconds and gives the answer. During the more slowly paced parts of the lesson, the teacher should wait for an answer as long as she feels that the child is thinking and will answer, but not so long as to embarrass the child or lose the other children's attention.

If the child does not respond within a reasonable time, the teacher should indicate that some response is expected by probing ("Do you know?"). She should then simplify (see #19) according to the type of question.

WHEN THE CHILD'S ANSWER IS INCORRECT

18. The teacher should indicate that the answer is wrong, and then follow simplification procedures outlined below for the two types of questions.

SIMPLIFICATION PROCEDURES

19. The appropriate simplification procedure is determined by the type of question.

a. If the question deals with factual knowledge that cannot be reasoned out, the teacher should give the answer to the child and then move on.

b. If the question is one that the child could reason out with help, the teacher should provide clues or simplify the question. If the clues still do not help the child, he should be given the answer. The teacher should never ask another child to supply the answer.

WHEN THE CHILD IS CORRECT

20. The teacher should acknowledge the correctness, and make sure that everyone else heard and understood the answer.

PRAISE AND CRITICISM

21. Praise is important but should not be used indiscriminately. Praise thinking and effort more than just getting the answer, and make praise as specific and individual as possible.

22. Criticism should also be as specific as possible and should include specification of desirable or correct alternatives.

Method

Nine schools and 27 female first grade teachers were involved in the study. Scores of 557 students were included in the analyses presented in this paper.

All data were collected in public elementary schools serving a predominately white population in Austin, Texas. All classrooms were traditional in their first-grade reading group instruction, in that

reading was taught in small ability groups in self-contained classrooms. Each of the nine schools was described by the school district as "middle socioeconomic status (SES)," although there was a small range of variation within this classification. Therefore, the nine schools were divided into three groups which were matched for SES and number of teachers. These three groups were then randomly assigned to the experimental groups. All participating teachers within any given school were in the same experimental group. This meant that school effects could interact with treatment effects, but it eliminated the problem of contamination of control teachers by experimental teachers.

The three groups were:

1. Treatment-observed. The 10 teachers in this group were instructed in the 22 principles and were observed teaching their reading groups approximately once a week for most of the year. (Student N = 192)

2. Treatment-unobserved. The seven teachers in this group also were instructed in the 22 principles, but they were not observed during the year. (Student N = 147)

3. Control-observed. The 10 teachers in this observer-presence group did not receive any specific treatment. They were told that the researchers were interested in learning about effective teaching behaviors for first grade reading instruction. They also were observed approximately once a week for most of the year. (Student N = 218)

The last two groups were included in the design to answer certain questions about the treatment. The treatment-unobserved group was included to determine if the treatment was effective without observation. The

control group was included to determine to what extent the principles were implemented naturally, without any treatment. Because the instructional model was developed by observing teachers and relating their behaviors to effectiveness, it was expected that most teachers would already be using these principles, in varying degrees. Also, by gathering observational data on the use of each principle, we would be able to look at process-product relationships across the two observed groups to determine each principle's relative effectiveness. That is, we could examine certain specific behaviors which were part of the large model (such as calling on the children in order rather than randomly) and assess their effects independently of the effects of the model as a whole.

When the teachers were contacted initially, they were told only about their assigned condition. Therefore, they were naive as to the inclusion of the other two groups for comparison. The control teachers were told only that we were interested in observing what they did in the classroom and relating it to achievement gains. They were given no suggestions about how to teach at any time.

The two treatment groups were given identical booklets which explained the 22 principles and the rationales behind them. They were encouraged to use the techniques in their teaching throughout the year, but were assured that they should use their own best judgment about how the techniques worked with their children. No further training with the principles was offered after the initial discussion, although the teachers kept the written materials for reference.

The 20 teachers who were observed (the treatment-observed and control groups) were seen approximately once a week. Their reading groups were

coded with an observational system designed for this study (Brophy, et al., Note 1). This system measured not only implementation of the 22 principles, but also the immediate effects of each of them. For example, if the teacher used a signal to call a group together, the classroom observer would note this, note what kind of signal was used, and note how effective it was in terms of time spent in transition and attention of the children to the signal. Other measures taken with the system will be used to evaluate the principles further. For example, was the frequency of call outs higher or lower in groups in which teachers used patterned turns rather than volunteers?

All students in the school system were given the Metropolitan Readiness Test at the beginning of the year. The scores were made available to us, and they served as covariants in the final analysis of gains. At the end of the year, the children in all of the classrooms involved in the study were tested with the Metropolitan Achievement Test, Primary I level. Due to the setting of the study and time limitations, only the three subtests pertaining to reading were given: word knowledge, word analysis, and reading. There also was a "total reading score" computed from the word knowledge and reading subtests.

In addition to the testing of the children and the classroom observations, other data collected were ratings of the teachers and students by the classroom observers and an end-of-the-year interview with each teacher, in which she rated her frequency of use of each of the principles and gave her opinion, of its general value and specific advantage and disadvantages.

Data Analyses

For the results reported in this paper, the measures of interest are the pretest and posttest scores of the children in each of the three treatment groups. Table 2 presents the mean scores for each group. The two measures were used to compute adjusted post scores for each of four posttest measures for each child, using the readiness scores as co-variables. A different set of predictors was used for each subtest of the posttest measure (the Metropolitan Achievement Test), because initial analysis revealed that certain subtests of the Metroplitan Readiness Test were better predictors of performance on certain of the subscores of the achievement test than were other readiness subtests. These analyses involved computing partial correlations between each readiness and achievement subtest. The readiness subtests with significant partial correlations with an achievement subtest were considered the best predictors for that measure. The combination of best predictors was different for each subtest of the Achievement Test, so results were computed separately for the four achievement subtests. The predictors used for each achievement subtest and the partial r's are listed in Table 1.

One-way analyses of variance of the adjusted achievement scores of the three groups were computed in two different ways, once using the 557 students' scores as the units of analysis, and once using the 27 teachers' class mean adjusted achievement scores as the units of analysis. Both types of analyses yielded similar results, but the analyses done with students' scores produced much higher significance values, due to the large sample size. Because the treatment was administered to the teachers rather than to the students (although separate implementation data are available

for each individual child observed) and because using teachers' scores as the units of analysis is more conservative, only the results for teachers will be reported here. However, the analyses of student scores served to strengthen the findings of the analyses of teachers' scores.

In addition to the three-group analyses of variance, post hoc group comparisons were done in order to determine the source of significant differences. This was done by means of multiple regression equations in which pairs of treatment groups were examined for their contribution to the prediction equation. That is, paired groups 1 and 2, 1 and 3, and 2 and 3 were analyzed by collapsing group membership vectors within a multiple regression equation; then these restricted models were each compared to a "full" model in which each treatment group was separately assigned to binary predictor vectors. This approach is described extensively in Ward and Jennings (1973). These analyses resulted in F-ratios which represented the difference in predictive accuracy between the full models and the restricted models (in which the pairs of groups were collapsed). A significant F indicated that the restricted model did not predict the criterion (the adjusted achievement scores) as well as the full model did. This meant that there was a significant difference between the two groups which were characterized by a single predictor vector in the restricted model. This approach is similar to a one-way, two-group analysis of variance in which the null hypothesis of equality between groups is tested, but by using regression equations, the contribution of all three groups to the total variance is taken into account in computing the F, so that the error term is reflective of total variance instead of only the within-group variance of the two groups in question.

Results

Analyses were done for each of the four posttest scores: word knowledge, word analysis, reading, and total reading. Due to the small sample size and predicted direction of differences, F values with $p \leq .10$ were considered significant. This is equivalent to $p \leq .05$ for one-tailed tests.

The one-way, three group analyses of variance indicated a significant difference between the three groups for the first three measures, and an almost significant effect for the total reading gain scores. The means and standard deviations of the adjusted post scores of the three groups are given in Table 3. The analysis of variance source tables are presented in Table 4.

The group comparisons yielded no significant differences on any measure between the two treatment groups (observed and unobserved). When the treatment-observed group and the control group were compared, significant differences were found on all measures except word analysis. Comparisons of the treatment-unobserved and the control groups resulted in significant differences for all four measures. The F ratios and p levels of the group comparisons are given in Table 5. All differences were in the predicted direction: treatment groups outperformed the control group.

Discussion

The results indicate that the treatment group teachers produced greater reading gains in their students than did the control group teachers, and that the treatment effect was not moderated by the presence of observers in the classroom.

Considering that the treatment itself was minimal in terms of training effort, these significant differences have immediate practical importance. They suggest that implications of research can be implemented in the classroom easily with measurable results.

It is tempting to present the instructional model as the cause of group differences, and to start recommending it as an effective, inexpensive way to improve instruction in first grade reading. However, experience in classrooms has convinced us that broad conclusions about what is the "best" approach to teaching are seldom, if ever, accurate and suitable for all children in a classroom. Therefore, rather than try to interpret the results at this point as due to a simple "treatment effect," we are presenting them as initial analyses which now may be examined in greater detail.

Further analyses will be performed using the observational data, the observers' ratings, and the teachers' interviews. The following questions will be addressed in order to define exactly what produced the differences in performance between groups:

(1) What are the relationships between learning and the specific parts of the instructional model? That is, do some teaching behaviors recommended by the model have more of an effect than others? For example, is it important to use sustaining feedback to errors but not important to use a standard signal for transitions or to use overviews?

(2) Are there aptitude-treatment interactions? That is, do some aspects of the instructional model play a more important role in promoting learning among lower-ability students than middle or high ability students? For example, is calling on children in order more important in low-level reading groups than in high-level reading groups?

(3) What are the effects of factors which were not specifically mentioned in the instructional model but which were measured by observation, rating, or interviewing? For example, were class size, reading group size, time spent in reading group, teachers' overall management ability, proportion of boys to girls, or number of seatwork assignments more predictive of reading gains than the treatment variables? Random assignment to treatment groups should have minimized differences between teachers in the three groups in these areas, but it is possible that there were systematic differences between the groups other than treatment assignment. Also, because all teachers within a school were assigned to the same treatment to prevent contamination, the possibility of strong school effects cannot be ruled out yet.

Answering such questions will yield more information about the classroom processes which are associated with learning gains in first-grade reading. Until these questions are examined, confident conclusions about the results would be premature. However, the data analyses reported here suggest that a strong treatment effect was present. If this is borne out by follow-up analyses, the data will have important implications for early reading instruction.

Reference Note

1. Brophy, J., Mahaffey, L, Greenhalgh, C., Ogden, J., & Selig, H. Coding system for the first grade reading group study. Report No. 75-2, Research and Development Center for Teacher Education, The University of Texas at Austin, 1975.

References

Brophy, J. & Evertson, C. Learning from teaching: A developmental perspective. Boston: Allyn and Bacon, 1976.

Southwest Educational Development Laboratory. Bilingual kindergarten program, inservice manual, Volume I. Austin, Texas: National Educational Laboratory Publishers, Inc., 1973.

Ward, J. & Jennings, E. Introduction to linear models. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973.

Table 1

Metropolitan Readiness Subtests Used to Compute

Predicted Metropolitan Achievement Test Scores

Metropolitan Achievement Subtest	Readiness Subtests Used as Predictors	Partial r's	p
Word Knowledge	Alphabet	.31	.001
	Numbers	.23	.001
Word Analysis	Alphabet	.30	.001
	Numbers	.24	.001
	Matching	.11	.005
Reading	Alphabet	.22	.001
	Numbers	.27	.001
	Listening	.11	.008
Total Reading	Alphabet	.27	.001
	Numbers	.27	.001
	Listening	.12	.002

Table 2

Means and Standard Deviations of Readiness and Achievement Subtest Scores

Metropolitan Readiness Tests: (raw scores)	Treatment-Observed		Treatment-Unobserved		Control	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Word Meaning	9.87	3.00	10.37	2.82	9.42	2.68
Listening	11.18	2.37	11.68	2.42	10.65	2.31
Matching	9.59	3.17	10.52	2.91	8.71	3.59
Alphabet	13.09	3.79	13.89	3.63	12.53	3.86
Numbers	14.99	5.07	16.47	4.72	14.88	4.73
Copying	8.22	3.82	8.45	3.27	9.41	3.53
Total Readiness	66.88	16.56	71.35	15.29	65.51	15.48
Metropolitan Achievement Tests (standard scores)						
Word Knowledge	48.29	11.42	51.06	12.14	45.21	10.05
Word Analyses	45.62	9.37	47.82	10.02	42.76	8.89
Reading	46.47	12.14	48.05	13.36	41.39	11.87
Total Reading	45.95	10.86	48.22	11.97	42.61	10.22

Table 3

Means and Standard Deviations of Adjusted

Post Scores of Three Groups of Teachers

		Treatment-Observed (N=10)	Treatment-Unobserved (N=7)	Control (N=10)
Word Knowledge	$\bar{X}$	1.04	2.27	-1.51
	SD	1.99	1.72	3.58
Word Analysis	$\bar{X}$	.56	1.78	-1.34
	SD	2.15	1.91	3.73
Reading	$\bar{X}$	1.67	1.38	-2.51
	SD	2.31	3.06	5.46
Total Reading	$\bar{X}$	.91	1.50	-1.59
	SD	2.19	1.88	4.41

Table 4

Results of Analysis of Variance of Adjusted

Post Scores of Three Groups of Teachers

Adjusted Achievement Score	Source	S.S.	df	M.S.	F	p
Word Knowledge	Treatment	65.00	2	32.50	4.62	.02
	Within groups	168.65	24	7.03		
Word Analysis	Treatment	42.46	2	21.23	2.70	.09
	Within groups	188.96	24	7.87		
Reading	Treatment	103.75	2	51.87	3.34	.05
	Within groups	372.89	24	15.54		
Total Reading	Treatment	48.81	2	24.40	2.45	.11
	Within groups	239.36	24	9.97		

Table 5

Results of Paired Group Comparisons

of Adjusted Post Scores

		Treatment-Observed vs. Treatment-Unobserved	Treatment-Observed vs. Control	Treatment-Unobserved vs. Control
Word Knowledge	F	.88	4.63	8.36
	p	.36	.04	.01
Word Analysis	F	.78	2.29	5.09
	p	.39	.14	.03
Reading	F	.02	5.60	3.99
	p	.88	.03	.06
Total Reading	F	.15	3.13	3.95
	p	.71	.09	.06